Journal of Law and the Biosciences, 747–753
doi:10.1093/jlb/lsv044
Response
Advance Access Publication 29 October 2015

Trust, vulnerable populations, and genetic data sharing

Jalayne J. Arias[1,*,†], Genevieve Pham-Kanter[2,‡], Rosa Gonzalez[1,**] and Eric G. Campbell[3,††]

1. Department of Bioethics, Cleveland Clinic
2. Department of Health Management and Policy, Drexel University School of Public Health
3. Massachusetts General Hospital and Harvard Medical School
*Corresponding author. E-mail: ariasj@ccf.org

ABSTRACT

Recent policies and proposed regulations, including the Notice of Proposed Rulemaking for the Common Rule and the 2014 NIH Genetic Data Sharing Policy, seek to improve research subject protections. Protections for subjects whose genetic data is shared are critical to reduce risks such as loss of confidentiality, stigma, and discrimination. In the article 'It depends whose data are being shared: considerations for genomic data sharing policies', Robinson et al. provide a response to our article, 'The Growth and Gaps of Genetic Data Sharing Policies'. Robinson et al. highlight the importance of individual and group preferences. In this article, we extend the conversation on models for improving protections which will mitigate

† Ms. Arias is the Associate Director of the NeuroEthics Program and Associate Professional Staff in the Department of Bioethics at the Cleveland Clinic. Ms. Arias' work incorporates empirical and conceptual projects addressing legal and ethical issues inherent in medicine and research.

‡ Dr. Pham-Kanter is an Assistant Professor in the Department of Health Management and Policy in the Drexel University School of Public Health. Her research focuses on policy questions related to physician–industry relationships and conflicts of interest in medicine. Other research interests include pharmaceutical and medical device policy, physician behavior and physician labor markets, empirical ethics and empirical health law, and health care quality and costs.

** Ms. Gonzalez is a research assistant in the Department of Bioethics at the Cleveland Clinic. Her work focuses on the ethical issues in preclinical biomarkers of Alzheimer's disease, genetic data sharing and traumatic brain injuries. She has Bachelor of Science in Biology from John Carroll University.

†† Dr. Campbell conducts research relating to physician conflict of interest and professionalism in medicine. Dr. Campbell is the Director of Research at Mongan Institute for Health Policy and a Professor at Harvard Medical School.

consequences for individuals and groups that are vulnerable to stigma and discrimination.

KEYWORDS: Data Sharing, genetic, genomic, research, policy, scientific norms

INTRODUCTION

The recently published Notice of Proposed Rulemaking (NPRM) for the Common Rule acknowledges the potential loss of public trust associated with high-profile disputes over biospecimens collected in research, including the Havasupai case and the Lacks Family settlement.[1] These events have consequences for the researchers involved and diminish benefits achieved through data sharing. The NPRM seeks to modernize the Common Rule and proposes two general changes affecting research with biospecimens. First, the NPRM redefines 'human subjects' to include biospecimens. Second, the NPRM revises consent requirements for research involving biospecimens. Specifically, the NPRM proposes a broad informed consent that would incorporate future research on the biospecimen collected from an individual. The NPRM, in addition to the recently issued National Institutes of Health Genetic Data Sharing Policy (GDS)[2], would alter human subjects' protections for genetic data sharing. While the NPRM and GDS both propose that these revisions will have a role in improving public trust through transparency, neither fully addresses ongoing challenges specific to trust and stigma for vulnerable individuals and groups in genetic data sharing.

In their response piece, 'It depends whose data are being shared: considerations for genomic data sharing policies', Robinson et al. evaluate human subject protections in genetic data-sharing policies and elaborate on the challenges inherent in balancing the benefit of data-sharing and protecting privacy.[3] The authors highlight the importance of public trust and provide the Global Alliance for Genomics and Health as a collaborative model for fostering trust. Research on individual preferences demonstrate that participants are generally supportive of data being shared among researchers.[4] However, as Robinson et al. point out, individuals and vulnerable populations may have differing views and preferences on how and under what circumstances data are shared.[5] Robinson et al. argue for research participant control over data and highlight dynamic consent as a potential model that would allow participants to specify and update their consent preferences.[6]

This article builds upon themes discussed by Robinson et al. and considers the role of consent and privacy protections in building trust with research participants, particularly

---

1 *Notice of Proposed Rulemaking, Federal Policy for the Protections of Human Subjects*, 80 Fed. Reg. 173 (proposed Sept. 8, 2015) (to be codified at 45 C.F.R pt. 46).

2 NIH Genomic Data Sharing Policy, NOT-OD-14–124GDS (Aug. 27, 2014).

3 Jill O. Robinson et al., *It Depends Whose Data Are Being Shared: Consideration for Genomic Data Sharing Policies*, J. L. & Biosci. 1, 8 (2015).

4 Susan Brown Trinidad et al., *Genomic Research and Wide Data Sharing: Views of Prospective Participants*, 12 Genet. Med. 486, 489 (2010); Robinsons, *supra* note 3; Amy McGuire et al., *To Share or Not Share: A Randomized Trial of Consent for Data Sharing in Genome Research*,13 Genet. Med. 948 (2011).

5 Robinson, *supra* note 3.

6 Jane Kaye et al., *Dynamic Consent: A Patient Interface for Twenty-first Century Research Networks*, 23 Euro. J. Hum. Genet. 141 (2015).

those who are members of a vulnerable population. In addition to considering the roles of individual consent, we also consider how vulnerable populations may be consulted to improve protections from stigma and discrimination.

WHAT IS A VULNERABLE POPULATION?

Vulnerable populations have previously been defined as 'social groups who have an increased relative risk or susceptibility to adverse health outcomes'.[7] Federal regulations provide specific protections to pregnant women, neonates, fetuses, prisoners, and children.[8] As a default, these groups have been characterized as 'vulnerable' populations for purposes of research protections. Yet neither definition of 'vulnerability' matches the vulnerability inherent in genetic data-sharing. Individuals and groups who are vulnerable for purposes of data-sharing may or may not also be groups, or members of groups, who are at increased risks of poor health outcomes. Regardless, it is not the potential health outcome which makes them vulnerable in this context. Instead, it is the potential loss of privacy and confidentiality which may expose individuals or populations to risks of stigma or discrimination based upon conclusions of research. These outcomes may not have been apparent to research participants when they provided consent.

Social groups categorized according to (i) race and ethnicity, (ii) disease population, and (iii) American Indians/Alaskan Natives lineage[9] may be particularly vulnerable to discrimination and stigma associated with genetic data sharing, based in part on a history of discrimination.[10] Discrimination may occur at two levels. First, at the individual level, disclosure of confidential information linking an individual's identity with genetic data connected to a stigmatized population may result in consequences for the individual. If research determines that an individual, who has consented to data being shared through a repository, is genetically disposed to develop Alzheimer's disease, the individual may be at risk of workplace discrimination or other forms of stigma. At the individual level, revision to consent procedures and improved privacy protections may mitigate some risks.

Second, at the group level, research that aims to identify an increased risk factor based on genetic status may have unanticipated consequence for a particular group. For example, the information gathered by Arizona State University which led to the case against it by the Havasupai tribe, indicated an increased risk for schizophrenia.[11] While this information may have some value for predicting health outcomes, it also placed significant stigma on the tribe.[12] Despite the history of incidents like the

---

7 Jacquelyn H. Flaskerud & Betty J. Winslow, *Conceptualizing Vulnerable Populations Health-Related Research*, 47 Nurs. Res. 69 (1998).

8 45 C.F.R. §§ 46.201, 301, 401.

9 American Indians/Alaskan Natives may also be considered within 'race and ethnicity", historically cultural beliefs and sovereignty issues inherent to American Indians and Alaskan Natives have created unique challenges for researchers and institutional review boards.

10 Jalayne J. Arias et al., *The Growth and Gaps of Genetic Data Sharing Policies in the United States*, 2 J. L. & Biosci. 56 (2015); see also Shawneequa Callier, et al., *Genetic Data-Sharing: What Will be Our Legacy?* 5 Front. Genet. Art. 34 (2014); William C. Rencher & Leslie E Wolf, *Redressing Past Wrongs: Changing the Common Rule to Increase Minority Voices in Research*, 103 Am. J. Pub. Health 2136 (2013).

11 Rosalina James et al., *Exploring Pathways to Trust: A Tribal Perspective on Data Sharing*, 16 Genet. Med. 820 (2014).

12 *Id.*

HeLa Cells release and the Havasupai case, neither the GDS nor the NPRM explicitly provides protections specific to vulnerable groups. Additionally, individual informed consent or privacy protections are not sufficient to reduce the risk of stigma associated with groups. It is not the release of private information alone,[13] but also the research results that place labels on a group which may expose members to stigma and discrimination.

Consent as a Starting Point: Individual Protections

The NPRM shifts how biospecimens are treated, from a discarded tissue sample described in *Moore v. Regents of the University of California* to an extension of the human subject.[14] As a result, the mode of consent and subject control of data must also shift. As Robinson et al. articulated, numerous repositories support genomic data and personal health information, including private repositories (ie, Google Genomics and Apple ResearchKit) and federally supported repositories (ie, Cancer Genomics Hub (CGHub) and dbGaP). The GDS provides guidance to researchers depositing into CGHub and dbGaP. However, unless a private institution receives NIH funding, the institution is not within the scope of the GDS.[15] As a result, private institutions' policies regarding research subjects' protections may vary. Diverse protection mechanisms and informed consent models are starting points for providing protections at the individual level and engendering trust.

Broad and dynamic consent models are proposed in different forms through policies and institutional recommendations. Broad consent, currently required by the GDS and proposed by the NPRM, seeks to inform participants that the specimen will be shared and used by future researchers.[16] Broad consent is currently used by several repositories, including the Alzheimer's Disease Neuroimaging Initiative and the Oxford BioBank.[17] This model of consent may inform participants and document their willingness to participate, but it fails to provide the level of individual control that is highlighted by Robinson et al. Upon consent, researchers and Data Access Committees (DACs) gain the power to determine what type of research may be appropriate, without requiring consideration of individual or group values. Additionally, it is unclear whether broad consent can provide sufficient information about future research that would allow a participant to fully understand the consequences of consent.

Proponents of the dynamic consent model propose to utilize technology to create a digital informed consent process.[18] Dynamic consent would allow participants to give or revoke consent to research involving their samples/information, provide a

[13] NPRM, *supra* note 1 at 54047 ('private information' is defined as 'information that has been provided for specific purposes by an individual and that cat the individual can reasonably expect will not be shared or made public').

[14] *Moore v. Regents of Cal.*, 793 P.2d 479 (1990).

[15] NIH GDS, *supra* note 2.

[16] Timothy Caulfield & Jane Kaye, *Broad Consent in Biobanking: Reflections on Seemingly Insurmountable Dilemmas*, 10 Med. Law Int'l 85 (2009)

[17] Oxford Biobank, Consent Form, http://www.oxfordbiobank.org.uk/wp-content/uploads/2013/11/OBB-consent-form-Version-1.1–25–9–2013.pdf (accessed Sept. 14, 2015); Alzheimer's Disease Neuroimaging Initiative, Introduction and Procedures for Accessing Data from Whole Genome Sequencing of ADNI Subjects, http://adni.loni.usc.edu/wp-content/uploads/2010/09/ADNI_WGS_Notice_20130917.pdf (accessed Sept. 14, 2015).

[18] Kaye, *supra* note 6.

virtual record of transactions, alter contact information, learn and consent to new research projects, and complete online surveys. Under the proposed model, a participant can adjust consent preferences to give broad consent or be contacted for each study. This allows the participant to be as engaged or removed from the process, according to their preferences, while maintaining control over their samples. This approach is consistent with research by McGuire et al. which reported participant preferences for a tiered decision-making structure.[19] The dynamic consent model may be consistent with literature reflecting variation of individuals and group preferences, as highlighted by Robinson et al.[20] However, dynamic consent precludes samples from being de-identified.[21] If an individual no longer wants to include their sample in research or decline a specific study, the sample must be re-identified and removed. As a result individuals are at an increased risk of loss of confidentiality. While the dynamic consent model is not currently implemented, the Apple ResearchKit has some features of the approach. Participants of the ResearchKit are able to select studies to participate in, information they would like to provide, and to leave a research study at any time.[22] Apple ResearchKit allows researchers to upload consent documents consistent with the researcher's institution IRB policies, add visuals to help participants understand key points, ask the participant if they would like to share data, and create a quiz to ensure the participants' knowledge of the study.

The Role of Data Access Committees

DACs serve a role in providing protections by regulating access to secondary researchers. The NIH DACs are charged with reviewing requests for access to data stored within the dbGaP and ensuring that requests are consistent with the data use limitations of the data set.[23] For data repositories not funded by federal funds, a range of approaches may be used to approve secondary research proposals. Protections provided by DACs are limited. Sufficient protections would require initial researchers to place appropriate limitations on data that restrict secondary use with consequences for a given vulnerable population. This also requires primary researchers anticipate how data may be used by secondary researchers. Secondly, DACS do not allow for the incorporation of individual or group preferences identified by Robinson et al. A DAC review does not provide any guarantee that data will not be used by secondary researchers for research purposes beyond an individual's or group's preference. This is particularly true for research participants who provide broad consent to research enrollment. Finally, enforcement for violation of the terms and conditions of the research agreement is not fully defined by the GDS, which simply states that 'the NIH will take appropriate action.' Insufficient enforcement and a lack of sensitivity to individual and group preferences reduce the potential for providing protections which may be critical to protecting vulnerable populations and engendering trust.

---

[19] McGuire, *supra* note 4.

[20] Robinsons, *supra* note 3.

[21] NPMR, *supra* note 1, at 53942–3 (distinguishing between non-identified, non-identifiable and de-identified. De-identified is distinguished as the HIPAA standard of non-identifiability).

[22] Apple ResearchKit, GitHub, *ResearchKit Framework Best Practices,* https://github.com/ResearchKit/ResearchKit/wiki/best-practices (accessed Sept. 10, 2015).

[23] National Institute of Health, Genetic Data Sharing, *NIH Data Access Committees and Chairs*, https://gds.nih.gov/04po2_1DAC.html (accessed Sept. 14, 2015).

Protections for Vulnerable Populations

Broad and dynamic consent models at the individual level are likely insufficient to reduce consequences for vulnerable populations at the group level. Broad consent limits the avenue for individuals or vulnerable groups to weigh in on research which may have consequences for the group (ie, linking a sickle cell trait to African Americans).[24] Vulnerable groups may be inadvertently harmed by the results of research utilizing broad consent where an individual's demographics are collected alongside the sample. However, there are prior examples of resolutions, policies, and research design methods which address challenges at the group level. The Lacks family settlement provided for a Lacks family member representative to be 'at the table' for consideration of research that would use the HeLa genetic line.[25] Similar arrangements have been developed for community representation in research design, particularly within tribal communities. For example, the Inter Tribal Counsel of Arizona has agreements between tribes and local universities to address issues (ie, protection of cultural interests) which may arise in research enrollment of tribe members.[26] This raises the question: Should group representatives should be included in the review of secondary researcher requests to access data? Should vulnerable groups have a representative who can agree or disagree on behalf of the group on whether to consent to research which may have consequences for vulnerable populations?

Incorporating individual and community perspectives in research development and execution is increasingly prevalent. The Patient Centered Outcome Research Institute requires that researchers include a member of the patient population in the design of research funded by the Institute.[27] The Food and Drug Administration includes patient representatives on its advisory committees and has begun considering methods further incorporating patient perspectives during risk-benefit assessment in the drug-approval process.[28] Similarly, Rencher and Wolf propose that the Common Rule require community consultation in the institutional review board process for research involving vulnerable populations.[29] And finally, community advisory boards in disease populations have been successful in informing research protocol development.[30]

Robinson et al. emphasize individual and group preferences. For example, they report on research demonstrating that HIV-positive individuals prioritize privacy over advancing research. While dynamic consent has been proposed as a method of respecting individual preferences, another method would be to include group representatives when reviewing requests for access to data by secondary researchers. Building upon the models discussed above, including group representatives as members of DACs or similar committees could provide an avenue for extending protections at the group level to mitigate potential stigma and discrimination. Currently, dbGaP DACs are made up of 'senior Federal employees with appropriate scientific, bioethics, and human subjects'

---

[24] Rencher, *supra* note 10.
[25] Callier, *supra* note 10.
[26] Rencher, *supra* note 10.
[27] Patient Centered Outcome Research Institute (PCORI), http://www.pcori.org/ (accessed Sept. 14, 2015).
[28] FDA statement: http://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM125651.pdf (accessed Sept. 14, 2015).
[29] Rencher, *supra* note 10.
[30] *Id.*

research expertise'.[31] Including representatives from vulnerable populations in the process would encourage DACs or other similar committees to be informed of and consider cultural or other group-centric issues prior to approving requests to access data. There are some limitations and challenges to this proposal. First, this would require identifying some criteria by which a research proposal includes or affects a vulnerable population. An initial step would be defining vulnerable populations for this purpose and whether the research has consequences for the population. While previously articulated definitions of 'vulnerability' may be informative, vulnerability for this context should reflect the nuance of risks associated with data-sharing. Second, this additional requirement would increase regulatory burden on investigators, possibly slowing down the research process or preventing some kinds of potentially helpful population-level research. It might also generate bias if not all observations from the study population are treated similarly, or limit the validity of research findings if some research tools cannot be used in a study.

Finally, we have drawn a distinction between protections at a group level and protections at an individual level. Given the heterogeneity in preferences at an individual level, there may well be tensions between group preferences and the preferences of an individual member within that group. A model for how these conflicts can best be resolved will need to be developed.

CONCLUSION

Issues around vulnerable populations, individuals' values, and preferences about data-sharing will continue to develop and shift. The Tuskegee Syphilis Study, Lacks Family settlement, and Havasupai case have shaped research and scientific policies and practices. Future events and advancements in science will continue to alter how human subject protections are evaluated and implemented. A recent survey of geneticists conducted by Pham-Kanter, Zinner, and Campbell found that 42 per cent of geneticists reported patient confidentiality as a very important or moderately important reason for withholding data, in contrast to 21 per cent just 10 years prior.[32] This survey also found that data repositories were perceived by scientists to be an important facilitator of data-sharing in genetics and the life sciences, and the use of repositories has dramatically expanded over the past decade. Negotiating the tensions inherent in these trends, and their implications for research participants and particularly vulnerable populations, will be an ongoing policy task.

The inclusion of biospecimens within the definition of 'research subjects' will create new requirements, including but not limited to consent for researchers. As a result, genetic data-sharing policies should incorporate mechanisms to guide how access and secondary use is regulated to mitigate stigmatizing populations vulnerable to discrimination. Consent policies and engaging group representatives in the DAC review process can be constructive additional steps to engage research participants, engender trust, and provide protections from stigma and discrimination.

---

[31] National Institutes of Health, *supra* note 23.

[32] Pham-Kanter, Genevieve, Darren E. Zinner, & Eric G. Campbell. 2014. *Codifying Collegiality: Recent Developments in Data Sharing Policy in the Life Sciences*. PLOS ONE 9(9):e108451. doi:10.1371/journal.pone.0108451.